Steps to Qualify or Validate Data after a Failed Critical Criteria Checks

In order to address issues related to the recent OIG Management Alert[1] associated with findings of failed 1-point quality control (QC) checks and data invalidation, EPA is providing some additional guidance on the process to validate or invalidate routine data based on an exceedance of important checks that have been identified as "critical criteria" in the QA Handbook[2]. These critical criteria checks[3] are part of a validation template that were developed for all criteria pollutants around 2006 by EPA and the monitoring organizations. Monitoring organizations, in their organizations specific quality assurance project plans may identify additional checks that they deem critical. The definition of the critical criteria can be found in Appendix D of the QA Handbook but the following quote is the driver behind this guidance:

> "Observations that do not meet each and every criterion on the Critical Criteria should be invalidated unless there are compelling reason and justification for not doing so."

Compelling evidence (reason) is data, such as (but not limited to) an independent audit point(s), a multi-point verification, or a prior zero/span check that establishes whether the analyzer was in fact operating within the percent difference critical criteria acceptance limits and whether the 1-point QC check itself is considered valid or invalid.

A valid QC check which exceeds acceptance criteria (i.e., "fails") will result in at least some data invalidation, but sometimes there is "compelling evidence" available regarding corrective actions and/or additional analyzer checks that may not be readily viewable in the AQS dataset that helps bracket the data set to be invalidated. A valid QC check is one that is conducted using certified, properly functioning equipment, conducted in a manner that adheres to appropriate procedures (SOPs). However, there may be cases (as described in scenarios #2 and #3 below) where additional information demonstrates that a QC check that exceeded the acceptance limits was for some reason, invalid. We need to use, evaluate and report both valid and invalid checks in a consistent manner.

The following three scenarios may exist for a monitor when a 1-point QC check has exceeded the established acceptance criteria. A flowchart follows that describes these three scenarios:

Scenario 1

1. A 1-point QC check exceeds the established acceptance criteria. Upon investigation, the operator determines that the 1-point QC check provided a valid concentration and that the analyzer needs adjustment/calibration. This confirmation provides evidence that the 1-Point QC check was a valid check and, resultantly, routine data should be invalidated.

Commented [JAT1]: I think there are circumstances such as high sensitivity analyzers with exceedingly low precision points in which a generalized justification for expanding the acceptable criteria is warranted.

Example:
A high sensitivity CO analyzer at a rural NCore site operates at a range of 5 ppm. CO concentrations rarely exceed 0.5 ppm and a precision level of 0.5 ppm is selected. Based on analyzer performance and specifications an acceptable zero drift of 0.1 ppm is determined as the control criteria. A check comes in and has a zero of 0.064 and a precision check of 0.568. The precision point is 11.4% off (greater than the 10% listed in the template) but it is clearly being driven by typical zero drift. In this case an organization may take the stance that an alternative criterial of +/- 0.1 ppm is appropriate and that the check and the data framed by the check is valid. The check still results in corrective action and recalibration as an action level has been exceeded per the organizations QAPP and SOPs.

[1] Report: Certain State, Local and Tribal Data Processing Practices Could Impact Suitability of Data for 8-Hour Ozone Air Quality Determinations [HYPERLINK "https://www.epa.gov/office-inspector-general/report-certain-state-local-and-tribal-data-processing-practices-could"]

[2] Quality Assurance Handbook for Air Pollution Measurement Systems Volume II Ambient Air Quality Monitoring Program [HYPERLINK "https://www3.epa.gov/ttn/amtic/qalist.html"]

[3] Although the guidance focuses on 1-point QC checks since it is the only check currently reported to AQS. There are other critical criteria that fall within the QA Handbook guidance

Flagging Process for Scenario 1

The 1-Point QC check is reported to AQS, and the null code "AS" (poor QA Results) replaces the routine data either back to the last acceptable 1-point QC check or where there is compelling evidence (i.e., an acceptable zero and span or other verification) to accept some of the data between the exceedance and the last 1-point QC check. In this case, the routine data that was valid would be reported and flagged "1V" while the data that was not supported by compelling evidence would be flagged with the null code "AS"

> **NOTE:** If no additional verification checks or other investigative measure to find compelling evidence is performed on the analyzer or the QC system following the QC exceedance, then the 1-point QC check will be considered valid and reported to AQS. EPA will consider the routine data suspect and the data should be replaced with the "AS" null code back to the last passing check and forward to the next passing check. Quarterly evaluation reports under development by EPA will highlight this data.

Scenario 2 and 3

2. A 1-point QC check exceeds the established acceptance criteria and **there is compelling evidence** to consider the analyzer's data valid. For example, after a failure the monitoring organization reviewed the data, went out to the site and conducted an "as is" (no adjustment to analyzer) QC check, performance evaluation, or multi-point verification at a concentration around the original QC check. These additional checks (not limited to the examples described above) demonstrate that the analyzer is operating within the 1-point QC acceptance limits and, therefore, supports the validity of the routine data. This compelling evidence also suggests that corrective action is needed to the QC system that generated the invalid 1-point QC check. It is suggested that corrective action be taken on the QC system immediately in order to determine the definitive cause of the invalid check, which serves as further evidence to support the validity of the routine data. A second acceptable 1-point QC check should be run so that routine data validity is established from the acceptable check to the next scheduled 1-point QC check.

3. Similar to scenario #2 where there is compelling evidence but a 1-point QC check was not run immediately after verifying that the analyzer is operating within acceptance limits, but was run within a few business days.

Flagging Process for Scenarios 2 and 3

The following process is for gaseous pollutant data that fail (exceed acceptance criteria) to meet 1-point QC checks (or Zero/Span) but monitoring organizations **have compelling evidence to consider the routine data valid** (scenarios #2 and 3). In other cases, where a monitoring organization responds to a failing QC check with an adjustment/recalibration, followed by a verification (ideally, followed by another QC check at the same concentration); the data after the multipoint calibration/ verification until the next passing p-check may be considered valid.

1. The invalid, failed 1-point QC check is not reported since the QC check is not considered valid. EPA will either create a flag field in the QA Transaction and will request the flag "IC" (invalid QC

Commented [JAT2]: The check in this case is valid, but the data is effectively removed. The result is that the precision and bias evaluations for the remaining data is still impacted by the failed check unless it is somehow excluded.

I think there is an argument for this approach that even though the directly impacted data has been removed that the check should stand as it is indicative of how the system is operating and other periods not bracketed by this check may still be subject to similar instability this check indicates.

Commented [JAT3]: It is very common for precision points that fail where the failure is driven by zero drift that the span will be perfectly fine. A worsening zero may provide evidence that it was the driving force and that prior to the zero drifting the data would have been valid.

Commented [JAT4]: WI typically has only reported "Pre" (As-Found/unadjusted/pre-maintenance) checks. An "Post" (As-Left/adjusted/post maintenance) check after a calibration adjustment is typically not reported. This would represent a change in philosophy for us in how we report checks and will increase the workload by 25-50% (roughly the amount of post adjustment checks we perform).

If possible I would like to see this document expanded to address EPA's stance on the reporting of Pre and Post checks explicitly.

Commented [JAT5]: This is unclear. If a second standard is brought in and the check passes, doesn't this bracket the data going back as well as the data going forward? What is the need for a "second acceptable 1-point QC check"?

Commented [JAT6]: Again, unclear. When we do an adjusted calibration we will always run a zero, span and precision point at a minimum after the instrument settings have been adjusted. In this case the 1-pt precision verification is part of our adjustment/recalibration routine. It then sounds as if a second QC check is ideally performed. Is this the same day? Why does this add value?

check) be reported in lieu of the QC value[4], or create an "IC" null code that can replace the 1-point QC concentration of the audit value (to be determined). This flag will create a "placeholder" in AQS that will allow to one identify that a QC check occurred within the required timeframe.

2. Routine data within the time frame between the last acceptable check and the next passing check should be flagged with a "1V" signifying the data was reviewed and there is a compelling reason to consider some or all of the data from the last passing QC check valid.
3. During the annual certification process, monitoring organizations will provide compelling evidence for the "1V" flags. The AMP600 Report will be modified to include ways for the monitoring organizations to provide the compelling evidence. As an option, monitoring organizations can provide free form comments in AQS. This comment can be entered via the web application on the maintain raw data form. EPA will work with monitoring organizations and provide additional guidance on this part of the process.
4. EPA Regions during the annual certification/concurrence process will concur with the data flagged "1V".

Commented [JAT7]: I am in favor of either mechanism. Small preference for the IC null code (leaves all data intact). This check should then be used in completeness metrics, but not in precision and bias metrics.

Commented [JAT8]: "compelling evidence" is very open to interpretation. Could some examples be provided?

The first two that come to mind are cross checks with alternative standards and graphical evidence that a point was taken before the analyzer had fully stabilized and was still approaching a valid response. This is common in automated checks that are compressed to prevent the loss of data (bridging the hour to leave hours >75% compelte).

Next Steps

Any routine data represented by failed 1-point QC checks that are not properly flagged in AQS will be identified in EPA quarterly evaluation reports (currently in design phase). EPA Regions will work with monitoring organizations on this data until a resolution of the validity of this data is reached prior to annual certification. Unresolved data issues represented by failed 1-point QC checks may not be considered for regulatory use until completion of steps 2-4.

In addition, 1-point QC checks will be evaluated for completeness in the quarterly reports to ensure a check is performed and reported (via a concentration or a flag) every 14 days. It is strongly suggested that these checks be automated to be performed more frequently than every 14 days to minimize loss of data due to invalidation. EPA Regions found monitoring organizations running checks more frequently but not reporting them to AQS. We suggest all valid QC checks be reported since it may also serve to minimize data invalidation.

EPA is in the review/development stages of this process. We will be working with the National Air Data Group to develop the flagging portion which can occur fairly quickly with the certification/concurrence part of the process to be ready before May 2018 annual certification and concurrence.

Commented [JAT9]: Convention for the every 14 days should be included here. I am currently opposed to the "first of the year" or "start of season" approach being used by AQS. This approach means for operators that in 2017 they need to get their checks done by a period ending on Saturday (perfect in my opinion) but by 2020 they will need to complete checks by the period ending on Tuesday. This moving due date makes for very difficult standardization of field operations.

14 day "buckets" should always be Sunday-Saturday regardless of what day the 1st of the year or ozone season falls on.

As organizations move to more automated checks this will be less of an issue, but we are not all there yet.

Commented [JAT10]: If an automated nightly precision check is done in additional to a manual biweekly this could result in 14 times more effort to report the checks. This will discourage the implementation of nightly checks of precision.

[4] At the time of this guidance AQS has not revised the QA Transaction to include a flag field. It is expected that this feature will be available starting in calendar year 2018.

Steps to Correctly Validate Data after a Failed Critical Criteria Checks